

相依误差下部分函数型线性模型的估计*

王亚飞

(北京工业大学应用数理学院, 北京 100124)

(E-mail: wangyafei@emails.bjut.edu.cn)

杜江 张忠占†

(北京工业大学应用数理学院, 首都社会建设和社会管理协同创新中心, 北京 100124)

(†E-mail: dujiang@bjut.edu.cn; zzhang@bjut.edu.cn)

摘要 本文研究当误差序列为平稳的 α -混合序列时, 部分函数型线性模型的估计问题, 基于用 Karhunen-Loève 展开来逼近斜率函数的思想, 给出了未知参数和斜率函数的估计方法, 并进一步建立了参数估计量的渐近正态性和斜率函数估计量的收敛速度. 最后用模拟研究和具体实例说明了估计方法的良好表现以及相依误差结构对估计量所带来的影响.

关键词 渐近正态性; α -混合序列; 主成分; 收敛速度

MR(2000) 主题分类 46B20

中图分类 O177.2

1 引言

为了对函数型预测变量和实值响应变量之关系进行建模, Ramasy 和 Silverman^[1] 提出了函数型线性模型. 随后, 许多学者对其进行研究, 例如, Cardot^[2], Cardot^[3], Cai 和 Hall^[4], Hall 和 Horowitz^[5], Li 和 Hsing^[6], Crambers^[7].

鉴于线性回归模型的适应性及函数型线性模型的已有方法论, Aneiros-Pérez 和 Vieu^[8] 提出了部分函数型线性模型:

$$Y = \boldsymbol{\beta} \mathbf{z} + \int_{\zeta} \gamma(t) X(t) dt + \varepsilon, \quad (1.1)$$

其中, $Y \in R$, \mathbf{z} 是一个 d - 维的随机变量且二阶矩有限, $X \in L^2(\zeta)$, $L^2(\zeta)$ 表示定义在

本文 2015 年 7 月 29 日收到. 2016 年 8 月 28 日收到修改稿.

* 国家自然科学基金 (11271039), 高等学校博士学科点专项科研基金 (20131103110027) 资助项目.

† 通讯作者.

概率空间 (Ω, \mathcal{F}, P) 上的均值为 0, 二阶矩有限的随机过程 (i.e. $E|X(t)|^2 < \infty, \forall t \in \zeta$), β 是 d - 维的未知参数向量, γ 是斜率函数且 $\gamma \in L^2(\zeta)$, ε 表示随机误差且独立于 \mathbf{z} 和 X . 对于部分函数型线性模型也已存在大量研究, 比如, Aneiros-Pérez 和 Vieu^[9] 使用非参数核回归的方法证明了参数分量估计量具有渐近正态性, 非参数分量估计量具有强相合性. 基于 Karhunen-Loève (K-L) 展开, Shin^[10] 利用最小二乘方法对该模型参数和非参数部分进行了估计, 并证明参数部分的估计量具有渐近正态性、斜率函数估计量具有最优的收敛速度. Shin 和 Lee^[11] 对部分函数型线性模型提出了岭估计方法, 并证明参数分量的岭估计量具有渐近正态性, 斜率函数的岭估计量与最小二乘估计量的收敛速度是一致的.

上述所提文献均假设误差 ε 是独立同分布 (i.i.d.) 的. 然而, 在现实生活中, 尤其是在经济领域, i.i.d. 假设并不成立. 鉴于实际需要性, 有必要考虑相依误差下部分函数型线性模型的估计. 本文从这个角度出发, 在误差序列形成平稳的 α - 混合序列时, 对部分函数型线性模型的估计问题进行了研究. 我们称随机变量序列 $\{\tau_i, i \geq 1\}$ 是 α - 混合的, 指当 $n \rightarrow \infty$ 时, 混合系数

$$\alpha(n) := \sup_{k \geq 1} \sup_{A \in \mathcal{F}_1^k, B \in \mathcal{F}_{n+k}^\infty} |P(AB) - P(A)P(B)|$$

趋于 0, 其中, $\mathcal{F}_n^m = \sigma(\tau_i, n \leq i \leq m)$ 表示由 τ_n, \dots, τ_m 生成的 σ - 代数.

在误差结构为 α - 混合序列时, 已有许多学者对普通的回归模型进行了研究. 例如, [12] 考虑了 α - 混合误差下的 E-V 模型, [13,14] 分别对含 α - 混合误差的非参数回归模型和纵向部分线性模型进行了研究. 在一定的正则条件下, 本文考虑了在误差结构是一个平稳的 α - 混合序列时, 模型 (1.1) 中未知参数的估计问题, 并研究了参数分量估计量的渐近性质和斜率函数估计量的收敛速度.

本文接下来的具体内容安排如下: 第 2 部分提出估计方法, 主要的理论结果在第 3 部分给出, 模拟研究和具体实例分别在第 4 和第 5 部分考虑, 附录给出了定理的证明及证明定理时所用到的预备引理.

2 估计

不失一般性, 假设 $\zeta = [0, 1]$. 空间 $L^2[0, 1]$ 上的内积和范数分别记为 $\langle \cdot, \cdot \rangle$ 和 $\|\cdot\|$. 过程 X 的自协方差函数记为 C_X 且假定是定义在 $[0, 1] \times [0, 1]$ 的连续函数, 利用 Mercer 定理^[15] 可知

$$C_X(s, t) = \sum_{j=1}^{\infty} \lambda_j \rho_j(s) \rho_j(t). \quad (2.1)$$

对应于 C_X 的协方差算子定义为

$$(T_X f)(\cdot) = \int_0^1 C_X(\cdot, t) f(t) dt.$$

式 (2.1) 中的 $\lambda_1, \lambda_2, \dots$ 和 ρ_1, ρ_2, \dots 分别表示核函数为 C_X 的协方差算子 T_X 的特征值和特征函数, 并且特征函数具有连续性和正交性. 为记号上的简便, 本文假设 $\lambda_1 > \lambda_2 > \lambda_3 > \dots > 0$. 由 K-L 展开可知

$$X(t) = \sum_{j=1}^{\infty} U_j \rho_j(t), \quad \gamma(t) = \sum_{j=1}^{\infty} \gamma_j \rho_j(t),$$

其中, U_j 是不相关的随机变量且 $EU_j = 0$, $E[U_j^2] = \lambda_j$. 利用上述展开, 模型 (1.1) 能够被写成

$$Y = \boldsymbol{\beta}' \mathbf{z} + \sum_{j=1}^{\infty} \gamma_j U_j + \varepsilon. \quad (2.2)$$

假设现有 n 组观测值 $\{(\mathbf{z}_1, X_1, Y_1), \dots, (\mathbf{z}_n, X_n, Y_n)\}$, 则模型 (2.2) 可写成

$$Y_i = \boldsymbol{\beta}' \mathbf{z}_i + \sum_{j=1}^{\infty} \gamma_j U_{ij} + \varepsilon_i, \quad i = 1, \dots, n, \quad (2.3)$$

其中, $U_{ij} = \langle X_i, \rho_j \rangle$, $\{\varepsilon_i, i = 1, \dots, n\}$ 形成一个平稳的 α -混合序列. 进一步我们可给出式 (2.3) 的逼近形式

$$Y_i \doteq \boldsymbol{\beta}' \mathbf{z}_i + \sum_{j=1}^m \gamma_j U_{ij} + \varepsilon_i, \quad i = 1, \dots, n, \quad m \rightarrow \infty. \quad (2.4)$$

然而, 在实际数据分析中, ρ_j 是未知的, 这样为得到估计量 $\hat{\boldsymbol{\beta}}$ 和 $\hat{\gamma}_j$, $j = 1, \dots, m$, 考虑使用 ρ_j 的估计形式 $\hat{\rho}_j$. 此时, 我们从式 (2.1) 的经验形式出发, 具体形式为

$$\hat{C}_X(s, t) = \frac{1}{n} \sum_{i=1}^n X_i(s) X_i(t) = \sum_{j=1}^{\infty} \hat{\lambda}_j \hat{\rho}_j(s) \hat{\rho}_j(t),$$

其中, $\hat{\lambda}_1 \geq \hat{\lambda}_2 \geq \dots \geq 0$, $(\hat{\lambda}_j, \hat{\rho}_j)$ 是核函数为 \hat{C}_X 的经验协方差算子的特征值和特征函数对. 令 $\mathbf{Y} = (Y_1, \dots, Y_n)'$, $\mathbf{Z} = (\mathbf{z}_1, \dots, \mathbf{z}_n)'$, $\mathbf{U}_m = (\langle X_i, \hat{\rho}_j \rangle)_{i=1, \dots, n, j=1, \dots, m}$, $\tilde{\boldsymbol{\gamma}} = (\gamma_1, \dots, \gamma_m)'$, $\boldsymbol{\varepsilon} = (\varepsilon_1, \dots, \varepsilon_n)'$. 根据式 (2.4), 模型 (2.2) 可被表示成

$$Y_i \doteq \mathbf{z}_i' \boldsymbol{\beta} + \sum_{j=1}^m \gamma_j \langle X_i, \hat{\rho}_j \rangle + \varepsilon_i, \quad i = 1, \dots, n,$$

用矩阵形式表示是

$$\mathbf{Y} \doteq \mathbf{Z} \boldsymbol{\beta} + \mathbf{U}_m \tilde{\boldsymbol{\gamma}} + \boldsymbol{\varepsilon}.$$

从而 $(\boldsymbol{\beta}, \tilde{\boldsymbol{\gamma}})$ 的估计 $(\hat{\boldsymbol{\beta}}, \hat{\tilde{\boldsymbol{\gamma}}})$ 可通过极小化函数

$$(\mathbf{Y} - \mathbf{Z} \boldsymbol{\beta} - \mathbf{U}_m \tilde{\boldsymbol{\gamma}})' (\mathbf{Y} - \mathbf{Z} \boldsymbol{\beta} - \mathbf{U}_m \tilde{\boldsymbol{\gamma}})$$

获得. 进一步地, 如果假设 $(\mathbf{Z}'(I - \mathbf{V}_m)\mathbf{Z})^{-1}$ 存在, 则 $\hat{\boldsymbol{\beta}}$ 和 $\hat{\tilde{\boldsymbol{\gamma}}}$ 的具体形式为

$$\hat{\boldsymbol{\beta}} = (\mathbf{Z}'(I - \mathbf{V}_m)\mathbf{Z})^{-1} \mathbf{Z}'(I - \mathbf{V}_m) \mathbf{Y},$$

其中, $\mathbf{V}_m = \mathbf{U}_m(\mathbf{U}_m' \mathbf{U}_m)^{-1} \mathbf{U}_m'$,

$$\hat{\boldsymbol{\gamma}} = (\mathbf{U}_m' \mathbf{U}_m)^{-1} \mathbf{U}_m' (\mathbf{Y} - \mathbf{Z} \hat{\boldsymbol{\beta}}).$$

进一步令 $\hat{C}_z = n^{-1} \sum_{i=1}^n \mathbf{z}_i \mathbf{z}_i'$, $\hat{C}_{zY} = n^{-1} \sum_{i=1}^n \mathbf{z}_i Y_i$, $\hat{C}_{zX}(t) = n^{-1} \sum_{i=1}^n \mathbf{z}_i X_i(t)$, $\hat{C}_{Xz}(t) = (\hat{C}_{zX}(t))'$, $\hat{C}_{YX}(t) = n^{-1} \sum_{i=1}^n Y_i X_i(t)$, 则 $\hat{\boldsymbol{\beta}}$ 、 $\hat{\boldsymbol{\gamma}}$ 存在如下等价形式

$$\hat{\boldsymbol{\beta}} = \left(\hat{C}_z - \sum_{j=1}^m \frac{\langle \hat{C}_{zX}, \hat{\rho}_j \rangle \langle \hat{C}_{Xz}, \hat{\rho}_j \rangle}{\hat{\lambda}_j} \right)^{-1} \left(\hat{C}_{zY} - \sum_{j=1}^m \frac{\langle \hat{C}_{zX}, \hat{\rho}_j \rangle \langle \hat{C}_{YX}, \hat{\rho}_j \rangle}{\hat{\lambda}_j} \right),$$

$$\hat{\gamma}_j = \langle \hat{C}_{YX} - \hat{\boldsymbol{\beta}}' \hat{C}_{zX}, \hat{\rho}_j \rangle / \hat{\lambda}_j, \quad j = 1, \dots, m,$$

其中,

$$\langle \hat{C}_{zX}, \hat{\rho}_j \rangle = n^{-1} \sum_{i=1}^n \mathbf{z}_i \langle X_i, \hat{\rho}_j \rangle, \quad \langle \hat{C}_{YX}, \hat{\rho}_j \rangle = n^{-1} \sum_{i=1}^n Y_i \langle X_i, \hat{\rho}_j \rangle.$$

相应地, 定义 $\gamma(\cdot)$ 的估计为

$$\hat{\gamma}(\cdot) = \sum_{j=1}^m \hat{\gamma}_j \hat{\rho}_j(\cdot).$$

3 渐近性质

假设观测值 $(\mathbf{z}_1, X_1), \dots, (\mathbf{z}_n, X_n)$ 是 i.i.d. 的, $\{\varepsilon_i\}$ 是平稳的 α -混合序列, 为得到所给估计量的渐近性质, 仍需如下假设条件. 令 C 表示正常数, 其值大小在不同地方可以不同.

(A1) 随机函数 X 平方可积且满足 $E\|X\|^4 < \infty$;

(A2) 存在常数 C 使得对任意的 j , $E[U_j^4] \leq C\lambda_j^2$;

(A3) 存在常数 C 和 $a > 1$ 使得 $C^{-1}j^{-a} \leq \lambda_j \leq Cj^{-a}$, $\lambda_j - \lambda_{j+1} \geq Cj^{-a-1}$, $j \geq 1$;

(A4) 存在常数 C 和 $b > a/2 + 1$ 使得 $|\gamma_j| \leq Cj^{-b}$, $j \geq 1$;

(A5) 对于截断项数 m , $m \sim n^{1/(a+2b)}$, 其中, $a_n \sim b_n$ 表示存在常数 $0 < L < M < \infty$ 使得对任意 n 有 $L \leq a_n/b_n \leq M$;

(A6) 对于随机向量 \mathbf{z} , $E\|\mathbf{z}\|_{R^d}^4 < \infty$, 其中, $\|\mathbf{z}\|_{R^d} = (\mathbf{z}'\mathbf{z})^{1/2}$; 此外, 对于 $j \geq 1$, $k = 1, \dots, d$, $|\langle C_{z_k X}, \rho_j \rangle| \leq cj^{-(a+b)}$;

(A7) 令 $f_k = \sum_{j=1}^{\infty} (\langle C_{z_k X}, \rho_j \rangle / \lambda_j) \rho_j$, $\eta_{ik} = z_{ik} - \langle f_k, X_i \rangle$, 则 $\eta_{1k}, \dots, \eta_{nk}$ 是 i.i.d. 的随机变量, 满足 $E[\eta_{1k} | X_1, \dots, X_n] = 0$, $E[\eta_{1k}^2 | X_1, \dots, X_n] = B_{kk}$, B_{kk} 是矩阵 \mathbf{B} 的第 k 个

对角元, 且 \mathbf{B} 是正定阵,

$$\mathbf{B} = E[\boldsymbol{\eta}_1 \boldsymbol{\eta}_1'] = C_{\mathbf{z}} - \sum_{j=1}^{\infty} \frac{\langle C_{\mathbf{z}X}, \rho_j \rangle \langle C_{X\mathbf{z}}, \rho_j \rangle}{\lambda_j},$$

其中, $\boldsymbol{\eta}_i = (\eta_{i1}, \dots, \eta_{id})'$, $C_{\mathbf{z}} = \text{Var}(\mathbf{z})$, $C_{\mathbf{z}X} = \text{Cov}(\mathbf{z}, X) = (\text{Cov}(z_1, X), \dots, \text{Cov}(z_d, X))'$, $C_{X\mathbf{z}} = C_{\mathbf{z}X}'$;

(A8) 随机误差满足 $E|\varepsilon_i|^{2+\delta} < \infty$ ($\delta > 0$), $E\varepsilon_i^2 = \sigma^2$, $E(\varepsilon_i) = 0$;

(A9) 存在 $r > 1 + 2/\delta$ 使得 $\alpha(n) = O(n^{-r})$;

(A10) 存在正整数 $p := p(n)$ 和 $q := q(n)$ 满足 $p + q \leq n$, $p/n \leq cq/p = o(1)$, 其中, p 和 q 分别是大块和小块的个数, 这将会在附录中进行具体介绍.

对于部分函数型线性模型, (A1)–(A7) 是标准的正则条件^[10,16]. 条件 (A8)–(A10) 是对相依误差下线性回归模型进行研究时常见的假设条件.

定理 3.1 若假设条件 (A1)–(A10) 成立, 则当 $n \rightarrow \infty$ 时, 有

$$n^{1/2}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) \xrightarrow{d} N(0, \sigma^2 \mathbf{B}^{-1}),$$

其中, “ \xrightarrow{d} ” 表示依分布收敛.

定理 3.2 在 (A1)–(A10) 假设条件下, 有

$$\|\hat{\gamma} - \gamma\|^2 = O_p(n^{-(2b+1)/(a+2b)}).$$

注 3.3 由定理 3.1, 3.2 的结论可知, 本文所给估计量的渐近性质与 [10] 所研究的独立误差结构下部分函数型线性模型估计量的渐近性质相似. 除此之外, 如果误差是 i.i.d. 的, 本文的理论结果与 [10] 中的结论一致. 另外需要补充说明的是, 一般情况下, 观测值 (\mathbf{z}_i, X_i) , $i = 1, \dots, n$ 是相关的, 此时 $\hat{\boldsymbol{\beta}}$ 的渐近协方差阵为 $\mathbf{B}^{-1} \Sigma_0 \mathbf{B}^{-1}$, 其中 Σ_0 是 Σ 的极限形式, 关于 Σ 的定义在定理的证明中给出. 另外也注意到, [10] 中的假设条件比本文要强.

注 3.4 为了应用所提方法, 我们需要知道怎样选择截断点 m . 理论上来说, 如果 m 取值过大, 模型 (2.4) 中的参数个数就会较多, 那么由函数型主成分分析 (FPCA) 的性质可知斜率函数 $\gamma(\cdot)$ 的估计可能就会变的很差; 如果 m 取值过小, 那么模型 (2.4) 就不能很好地逼近模型 (2.3), 此时可看出条件 (A5) 所起到的作用. 对于具体怎样选取 m , 已存在一些准则, 比如, AIC, BIC, 广义交叉验证 (GCV) 和 FPCA. AIC、BIC 和 GCV 是数据驱动准则, 它们能够在回归模型的拟合优度和复杂度之间进行平衡; FPCA 是从方差占比的角度进行考虑, 即前 m 个特征值占总方差的比例, 其中比例是事先给定的. 在模拟部分, 我们将分别考虑用 GCV 和 FPCA 准则选取 m .

注 3.5 可用 $\hat{\mathbf{B}} = (\mathbf{Z}'(I - \mathbf{V}_m)\mathbf{Z})/n$ 作为 \mathbf{B} 的估计, 并且附录中的式 (7.2) 保证了 $\hat{\mathbf{B}}$ 的相合性. 然而, 如果不知道 $\{\varepsilon_i\}$ 的具体协方差结构, 很难给出 σ^2 的相合估计. 第 5 部分将对这一点进行说明.

4 模拟研究

本小节我们通过 Monte Carlo 模拟研究来说明所提方法的有限样本性质. 样本数据由如下模型产生:

$$Y_i = z_{i1}\beta_1 + z_{i2}\beta_2 + \int_0^1 \gamma(t)X_i(t) dt + \varepsilon_i,$$

其中, $z_{i1} \sim N(0, 1)$, $P(z_{i2} = 0) = P(z_{i2} = 1) = 0.5$, $\gamma(t) = \sqrt{2}\sin(\pi t/2) + 3\sqrt{2}\sin(3\pi t/2)$. 由

$$X(t) = \sum_{j=1}^{200} U_j \rho_j(t), \quad t \in [0, 1],$$

产生样本函数 X , 其中, $U_j \sim N(0, ((j-0.5)\pi)^{-2})$, $\rho_j(t) = \sqrt{2}\sin((j-0.5)\pi t)$. 随机误差 ε_i 服从 AR(1) 模型:

$$\varepsilon_i = \tau\varepsilon_{i-1} + e_i,$$

其中, τ 的取值为 0.25, 0.5, 0.75, $e_i \sim N(0, 1)$. 为了与独立误差结构下模型估计作比较, 模拟中还考虑了 $\tau = 0$ 的情形. 对 τ 的每个取值, 分别考虑用 GCV 和 FPCA 准则选取 m , 样本量大小为 100, 300, 500, 1000, 且对每个样本量, 模拟 1000 次. 对于用 FPCA 选取 m , $m = \min \left\{ k : \sum_{i=1}^k \hat{\lambda}_i / \sum_{i=1}^n \hat{\lambda}_i \geq 0.85 \right\}$.

用表达式

$$\text{SRAE} = \frac{1}{1000} \sum_{i=1}^{1000} \left\{ \frac{1}{N} \sum_{s=1}^N [\hat{\gamma}_i(t_s) - \gamma(t_s)]^2 \right\}$$

评价斜率函数估计量的好坏, 其中 $t_s, s = 1, \dots, N$ 是等间隔的格子点, 本节中 $N = 100$, 下角标 i 表示第 i 次重复. 具体模拟结果见表 1 和表 2.

由表 1, 表 2 可得如下结论:

- 1) 固定 τ 的一个取值, MSE 和 SRAE 都随样本量的增大而减小, 且对模拟中所考虑的 τ 的不同取值都如此, 这表明本文所提方法的合理性.
- 2) τ 值越大表示误差序列的相关性越强, 在样本量固定时, MSE 和 SRAE 随 τ 的增大而增加.
- 3) 当样本量固定, τ 取 0 时, 也就是说误差独立, 此时 MSE 和 SRAE 的值要比相依情形下的值小, 这表明相依误差结构对估计量造成一定影响.
- 4) 对于用两种不同的准则选取 m , β 的估计效果相差不大.

针对模拟中所考虑的情形, 就 SRAE 而言, 如果用 FPCA 选取 m , $\gamma(\cdot)$ 的估计效果较好. 然而, 进一步的研究表明, 如果把方差占比提高, SRAE 的值也会增大. 比如, 当 $n = 100$, $\tau = 0.25$ 时, 把方差占比从 0.85 提高到 0.95, 则 SRAE 值变为 2.667, 比表 1 中的值 1.700 要大, 但是 β 估计量的偏差和 MSE 值不变.

模拟研究表明本文所提方法的可行性. 除此之外, 如果误差序列的协方差结构已知, 我们期望用加权最小二乘估计来提高估计量的有限样本性质. 关于这一点, 超出了本文的范围, 在此不作详细讨论.

表 1 GCV 准则下的有限样本性质

n	τ	β_1		β_2		γ
		Bias	MSE	Bias	MSE	SRAE
$n = 100$	0.00	0.005	0.103	0.000	0.144	1.588
	0.25	0.005	0.105	0.001	0.170	1.700
	0.50	-0.001	0.117	0.003	0.236	2.333
	0.75	0.002	0.149	-0.016	0.411	4.249
$n = 300$	0.00	0.001	0.060	0.001	0.084	0.592
	0.25	0.001	0.059	0.004	0.100	0.683
	0.50	0.000	0.069	0.005	0.133	0.795
	0.75	-0.002	0.089	0.008	0.240	1.303
$n = 500$	0.00	0.001	0.045	0.002	0.065	0.380
	0.25	0.000	0.046	0.002	0.078	0.383
	0.50	-0.002	0.052	-0.003	0.107	0.483
	0.75	-0.003	0.070	0.003	0.188	0.810
$n = 1000$	0.00	0.000	0.032	0.000	0.046	0.188
	0.25	-0.001	0.034	0.003	0.053	0.202
	0.50	-0.001	0.036	0.000	0.073	0.233
	0.75	0.000	0.046	0.001	0.136	0.381

表 2 FPCA 准则下的有限样本性质

n	τ	β_1		β_2		γ
		Bias	MSE	Bias	MSE	SRAE
$n = 100$	0.00	-0.003	0.103	0.001	0.149	0.768
	0.25	0.004	0.107	0.001	0.171	0.945
	0.50	-0.002	0.117	0.004	0.236	0.840
	0.75	0.001	0.148	-0.017	0.411	1.200
$n = 300$	0.00	0.000	0.058	0.003	0.084	0.145
	0.25	0.001	0.062	0.002	0.098	0.153
	0.50	-0.003	0.068	0.005	0.131	0.169
	0.75	0.002	0.084	0.003	0.257	0.248
$n = 500$	0.00	0.001	0.048	0.000	0.066	0.079
	0.25	0.000	0.046	0.002	0.078	0.081
	0.50	-0.002	0.052	-0.003	0.107	0.096
	0.75	-0.003	0.070	0.003	0.188	0.151
$n = 1000$	0.00	-0.000	0.033	0.002	0.045	0.039
	0.25	0.000	0.033	0.000	0.054	0.043
	0.50	-0.001	0.037	0.005	0.073	0.046
	0.75	-0.002	0.047	-0.001	0.135	0.073

5 实例分析

现收集到 1972 年 1 月到 2005 年 1 月 (397 个月) 的商业月用电量数据 (C) 和每年的平均零售价格 (P). 本小节的主要目的是利用本文所提方法对该数据集进行分析, 研究相依误差结构对 $\hat{\beta}$ 渐近方差的影响.

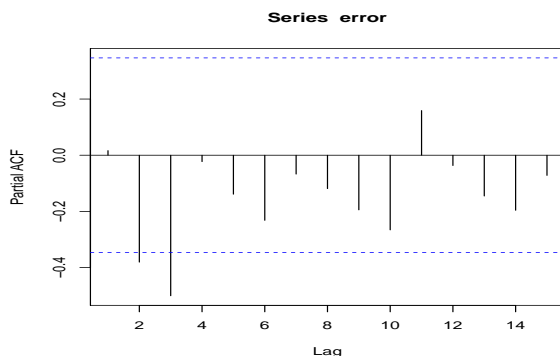


图 1 误差序列的偏相关函数

首先对月用电量数据集是否是平稳序列进行检验, 发现该数据集存在异方差性及线性趋势, 所以在对数据进行分析前进行对数差分. 具体做法如下:

$$D_j = \ln C_j - \ln C_{j-1}, \quad j = 1, \dots, 397,$$

$$X_i = \{D_{12(i-1)+t}, t \in [1, 12]\}, \quad i = 1, \dots, 32.$$

响应变量

$$Y_i = D_{12i+6}, \quad i = 1, \dots, 32,$$

模型中的实值变量定义为

$$z_i = P_i, i = 1, \dots, 32.$$

Y 对 Z 和 X 作回归可得残差序列 $\{\hat{\varepsilon}_i, i = 1, \dots, n\}$, 发现误差序列并不是 i.i.d. 的. 接下来探讨误差序列的具体相关结构, 首先利用 R 里面的程序包 “adf.test()” 对序列的平稳性进行检验, p 值是 0.19, 也就是说误差序列是平稳序列, 接着计算了该序列的自相关函数和偏相关函数. 如图 1 所示, 误差序列的相依结构可表示成:

$$\hat{\varepsilon}_i = \alpha_1 \hat{\varepsilon}_{i-2} + \alpha_2 \hat{\varepsilon}_{i-3} + e_i, \quad e_i \sim N(0, \delta^2), \quad i = 4, \dots, 32. \quad (5.1)$$

在给出 α_1, α_2 的估计后, 用 Ljung-Box 检验统计量对 $\{e_i, i = 1, \dots, n\}$ 进行独立性检验, p 值是 0.46, 表明序列 e_i 是独立的. 基于模型 (5.1), 计算可得

$$\text{Var}(\hat{\varepsilon}_i) = 2.30, \quad \text{sd}(\hat{\beta}) = 0.15.$$

如果忽略了序列 $\{\hat{\varepsilon}_i, i = 1, \dots, n\}$ 间的相依性, 可算得

$$\text{sd}(\hat{\beta}) = 0.18,$$

与相依情形相比变化了 20%, 间接说明了考虑相依结构的必要性. 最后需要补充说明的是, 在对该数据集分析时, 采用 FPCA 准则选取 m , 方差占比取 0.85.

注 5.1 如果想用该数据集预测未来用电量或者是基于模型 (1.1) 给出 β 的置信区间, 那么本节目标的重要性显而易见. 另外需要说明的是, 我们用 6 月份的用电量作为响应变量, 同样地, 也可以使用其它月份的用电量作为响应变量.

6 总结与讨论

对于部分函数型线性模型, 若误差序列是平稳的 α -混合序列, 本文提出相应的估计方法. 正如第 5 部分所示, 如果知道误差序列的具体相依结构, 那么基于所提方法可进行精确的统计推断. 由于平稳的 α -混合序列包括一大类相依结构, 所以本文方法具有较大的实用价值.

第 4 部分, 我们用 GCV 和 FPCA 准则选取 m , 并且在每种准则下给出了 SRAE 的值. 在模拟研究中, 我们同样考虑了用 AIC 和 BIC 准则选取 m , 鉴于篇幅所限, 没有列出具体结果. 通过模拟研究发现, 不论用哪个准则选取 m , 当样本量适中时, 对 β 的 MSE 值影响不大. 另外, 通过计算发现, 在选取 m 时, GCV 和 AIC 准则类似, 且选取的 m 值比用 FPCA 和 BIC 准则选取的 m 值波动大. 当样本量较小时, 为保证模型的拟合优度, 倾向于使用 GCV 或 AIC 准则, 但此时也要注意过拟合的情形. 如果用 FPCA 选 m , 正如我们所了解到的, 关于怎样做到在模型拟合优度和斜率函数估计量精度之间的平衡, 目前为止还不存在理论研究, 从而要注意方差占比的设定问题.

7 附录

设 $\{V_i, i \geq 1\}$ 是一个平稳的 α -混合随机变量序列, 混合系数记为 $\{\alpha(n)\}$.

引理 7.1^[17] 设 V_1, \dots, V_n 是一个 α -混合随机变量, 相应的 σ -代数记为 $\mathfrak{F}_{i_1}^{j_1}, \dots, \mathfrak{F}_{i_m}^{j_m}$, 且 $1 \leq i_1 < j_1 < \dots < j_m \leq n$, 对任意的 $l, j = 1, 2, \dots, m$, $i_{l+1} - j_l \geq \omega \geq 1$, $|V_j| \leq 1$, 有

$$\left| E\left(\prod_{j=1}^m V_j\right) - \prod_{j=1}^m EV_j \right| \leq 16(m-1)\alpha(\omega),$$

其中, $\mathfrak{F}_a^b = \sigma\{V_i, a \leq i \leq b\}$, $\alpha(n)$ 是混合系数.

引理 7.2 [18, 推论 A.2, 278 页] 假设随机变量 X 和 Y 满足 $E|X|^p < \infty$, $E|Y|^q < \infty$, 其中, $p, q > 1, p^{-1} + q^{-1} < 1$, 则

$$|EXY - EXEY| \leq 8\|X\|_p\|Y\|_q \left\{ \sup_{A \in \sigma(X), B \in \sigma(Y)} |P(AB) - P(A)P(B)| \right\}^{1-p^{-1}-q^{-1}}.$$

引理 7.3 ([19, 定理 4.1]) 设 $2 < p < q \leq \infty$, $2 < k \leq q$, $EV_n = 0$. 假设存在 $C > 0$, $r > 0$ 使得 $\alpha(n) \leq Cn^{-r}$ 成立. 如果 $r > k/(k-2)$, $r \geq (p-1)q/(q-p)$, 那么对 $\varepsilon > 0$, 存在 $Q = Q(\varepsilon, p, q, k, r, C) < \infty$, 使得

$$E \left| \sum_{i=1}^n V_i \right|^p \leq Q \left(n^{p/2} \max_{1 \leq i \leq n} \|V_i\|_k^p + n^{1+\varepsilon} \max_{1 \leq i \leq n} \|V_i\|_q^p \right).$$

引理 7.4 假设 (A1), (A8) 成立, 那么

$$\left\| n^{-1} \sum_{i=1}^n X_i \varepsilon_i \right\| = O_p(n^{-1/2}).$$

证 注意到

$$\begin{aligned} E \left\| n^{-1} \sum_{i=1}^n X_i \varepsilon_i \right\|^2 &= n^{-2} E \left\langle \sum_{i=1}^n X_i \varepsilon_i, \sum_{j=1}^n X_j \varepsilon_j \right\rangle = n^{-2} \sum_{i=1}^n \sum_{j=1}^n E \langle X_i, X_j \rangle E(\varepsilon_i \varepsilon_j) \\ &= n^{-2} \sum_{i=1}^n E \|X\|^2 E \varepsilon_i^2 = n^{-2} E \|X\|^2 \sum_{i=1}^n E \varepsilon_i^2 \\ &\leq cn^{-1} E \|X\|^2 = O(n^{-1}). \end{aligned}$$

所以有, $\|n^{-1} \sum_{i=1}^n X_i \varepsilon_i\| = O_p(n^{-1/2})$. 证毕.

定理 3.1 的证明 令

$$\begin{aligned} \widehat{\Phi}_k(f) &= \sum_{j=1}^m (\langle \widehat{C}_{z_k X}, \widehat{\rho}_j \rangle / \widehat{\lambda}_j) \langle \widehat{\rho}_j, f \rangle, \\ \Phi_k(f) &= \sum_{j=1}^{\infty} (\langle C_{z_k X}, \rho_j \rangle / \lambda_j) \langle \rho_j, f \rangle, \end{aligned}$$

其中, $f \in L^2[0, 1]$. 对 $A = (A_{ij}) \in R^{d \times d}$, 设 $\|A\|_{\infty} = \max_i \sum_j |A_{ij}|$. 注意到

$$\begin{aligned} n^{1/2}(\widehat{\beta} - \beta) &= \widehat{\mathbf{B}}^{-1} n^{1/2} \left\{ \frac{1}{n} \sum_{i=1}^n \left(\mathbf{z}_i - \sum_{j=1}^m \frac{\langle \widehat{C}_{zX}, \widehat{\rho}_j \rangle \langle X_i, \widehat{\rho}_j \rangle}{\widehat{\lambda}_j} \right) (\langle \gamma, X_i \rangle + \varepsilon_i) \right\} \\ &= \widehat{\mathbf{B}}^{-1} n^{-1/2} \left\{ \sum_{i=1}^n \left(\mathbf{z}_i - \sum_{j=1}^m \frac{\langle \widehat{C}_{zX}, \widehat{\rho}_j \rangle \langle X_i, \widehat{\rho}_j \rangle}{\widehat{\lambda}_j} \right) \langle \gamma, X_i \rangle \right. \\ &\quad \left. + \sum_{i=1}^n \left(\sum_{j=1}^{\infty} \frac{\langle C_{zX}, \rho_j \rangle \langle X_i, \rho_j \rangle}{\lambda_j} - \sum_{j=1}^m \frac{\langle \widehat{C}_{zX}, \widehat{\rho}_j \rangle \langle X_i, \widehat{\rho}_j \rangle}{\widehat{\lambda}_j} \right) \varepsilon_i \right. \\ &\quad \left. + \sum_{i=1}^n \left(\mathbf{z}_i - \sum_{j=1}^{\infty} \frac{\langle C_{zX}, \rho_j \rangle \langle X_i, \rho_j \rangle}{\lambda_j} \right) \varepsilon_i \right\}, \end{aligned}$$

其中, $\widehat{\mathbf{B}} = \widehat{C}_z - \{\widehat{\Phi}_k(\widehat{C}_{z_l X})\}_{k,l=1,\dots,d} = (\mathbf{Z}'(I - \mathbf{V}_m)\mathbf{Z})/n$.

根据引理 7.4, 类似于 [10] 中的证明可得

$$\|\widehat{\mathbf{B}} - \mathbf{B}\|_\infty = O_p(n^{-(2b-1)/(a+2b)}), \quad (7.1)$$

$$n^{-1/2} \sum_{i=1}^n \left(\mathbf{z}_i - \sum_{j=1}^m \frac{\langle \widehat{C}_{\mathbf{z}X}, \widehat{\rho}_j \rangle \langle X_i, \widehat{\rho}_j \rangle}{\widehat{\lambda}_j} \right) \langle \gamma, X_i \rangle = o_p(1), \quad (7.2)$$

$$n^{-1/2} \sum_{i=1}^n \left(\sum_{j=1}^\infty \frac{\langle C_{\mathbf{z}X}, \rho_j \rangle \langle X_i, \rho_j \rangle}{\lambda_j} - \sum_{j=1}^m \frac{\langle \widehat{C}_{\mathbf{z}X}, \widehat{\rho}_j \rangle \langle X_i, \widehat{\rho}_j \rangle}{\widehat{\lambda}_j} \right) \varepsilon_i = o_p(1). \quad (7.3)$$

接下来考虑项

$$n^{-1/2} \sum_{i=1}^n \left(\mathbf{z}_i - \sum_{j=1}^\infty \frac{\langle C_{\mathbf{z}X}, \rho_j \rangle \langle X_i, \rho_j \rangle}{\lambda_j} \right) \varepsilon_i.$$

令

$$\Sigma = \text{Var} \left(n^{-1/2} \sum_{i=1}^n \left(\mathbf{z}_i - \sum_{j=1}^\infty \frac{\langle C_{\mathbf{z}X}, \rho_j \rangle \langle X_i, \rho_j \rangle}{\lambda_j} \right) \varepsilon_i \right).$$

将证明如下结论

$$n^{-1/2} \Sigma^{-1/2} \sum_{i=1}^n \left(\mathbf{z}_i - \sum_{j=1}^\infty \frac{\langle C_{\mathbf{z}X}, \rho_j \rangle \langle X_i, \rho_j \rangle}{\lambda_j} \right) \varepsilon_i \xrightarrow{d} N(0, 1),$$

也就是说,

$$n^{-1/2} \Sigma^{-1/2} \sum_{i=1}^n \boldsymbol{\eta}_i \varepsilon_i \xrightarrow{d} N(0, 1), \quad (7.4)$$

其中, “1” 表示单位阵.

设 $w := w_n = \lfloor \frac{n}{p+q} \rfloor$. 我们使用 Bernstein 大块和小块技术^[12] 证明式 (7.4), 其中, p 和 q 分别表示大块和小块的个数. 记 $\mathbf{V}_i = \boldsymbol{\eta}_i \varepsilon_i / \sqrt{n}$, $\mathbf{A}_i = \boldsymbol{\eta}_i \varepsilon_i$,

$$\mathbf{V}_{mn} = \sum_{k=k_m}^{k_m+p-1} \mathbf{V}_k, \quad \mathbf{V}'_{mn} = \sum_{l=l_m}^{l_m+q-1} \mathbf{V}_l, \quad \mathbf{V}''_{wn} = \sum_{k=w(p+q)+1}^n \mathbf{V}_k,$$

其中, $k_m = (m-1)(p+q) + 1$, $l_m = (m-1)(p+q) + p + 1$, $m = 1, \dots, w$, 则,

$$\begin{aligned} \sum_{i=1}^n \mathbf{V}_i &= \sum_{m=1}^w \sum_{k=k_m}^{k_m+p-1} \mathbf{V}_k + \sum_{m=1}^w \sum_{l=l_m}^{l_m+q-1} \mathbf{V}_l + \sum_{k=w(p+q)+1}^n \mathbf{V}_k \\ &= \sum_{m=1}^w \mathbf{V}_{mn} + \sum_{m=1}^w \mathbf{V}'_{mn} + \mathbf{V}''_{wn} \\ &:= S'_n + S''_n + S'''_n. \end{aligned}$$

对于式 (7.4) 的证明, 只需要证

$$\mathbb{E} \|S''_n\|^2 \rightarrow 0, \quad \mathbb{E} \|S'''_n\|^2 \rightarrow 0, \quad (7.5)$$

$$\text{Var}(S'_n) = \Sigma + o(1), \quad (7.6)$$

$$\left\| \text{Eexp}\left(it \sum_{m=1}^w \mathbf{V}_{mn}\right) - \prod_{m=1}^w \text{Eexp}(it\mathbf{V}_{mn}) \right\| \rightarrow 0, \quad (7.7)$$

$$\text{E} \sum_{m=1}^w \|\mathbf{V}_{mn}\|^2 I_{(\|\mathbf{V}_{mn}\| > \varepsilon)} \rightarrow 0. \quad (7.8)$$

首先证明式 (7.5). 注意到

$$\begin{aligned} \text{E}\|S''_n\|^2 &= \sum_{m=1}^w \sum_{l=l_m}^{l_m+q-1} \text{E}\|\mathbf{V}_l\|^2 + 2 \sum_{m=1}^w \sum_{l_m \leq i < j \leq l_m+q-1} \text{Cov}(\mathbf{V}_i, \mathbf{V}_j) \\ &\quad + 2 \sum_{1 \leq i < j \leq w} \text{Cov}(\mathbf{V}'_{in}, \mathbf{V}'_{jn}) \\ &:= J_{1n} + J_{2n} + J_{3n}. \end{aligned}$$

由 (A8)–(A10) 及引理 7.2, 可证得 $J_{1n} = o(1)$, $J_{2n} = o(1)$, $J_{3n} = o(1)$. 从而式 (7.5) 中的第一个式子成立. 具体推导过程如下:

$$\begin{aligned} J_{1n} &= n^{-1} \sum_{m=1}^w \sum_{l=l_m}^{l_m+q-1} \text{E}\|\boldsymbol{\eta}_l \varepsilon_l\|^2 \leq n^{-1} \sum_{m=1}^w \sum_{l=l_m}^{l_m+q-1} \text{E}\|\boldsymbol{\eta}_l\|^2 \text{E}\varepsilon_l^2 \\ &\leq cn^{-1} \sigma^2 wq \leq c \frac{q}{p} = o(1), \\ \|J_{2n}\| &\leq n^{-1} \sum_{m=1}^w \sum_{l_m \leq i < j \leq l_m+q-1} \|\text{Cov}(\mathbf{A}_i, \mathbf{A}_j)\| \\ &\leq n^{-1} \sum_{m=1}^w \sum_{l_m \leq i < j \leq l_m+q-1} \alpha^{\frac{\delta}{2+\delta}} (j-i) \\ &\leq cn^{-1} wq \sum_{t=1}^{\infty} \alpha^{\frac{\delta}{2+\delta}} (t) \leq cn^{-1} wq = o(1), \\ \|J_{3n}\| &\leq n^{-1} \sum_{1 \leq i < j \leq w} \sum_{l=l_i}^{l_i+q-1} \sum_{k=l_j}^{l_j+q-1} \|\text{Cov}(\mathbf{A}_l, \mathbf{A}_k)\| \\ &\leq cn^{-1} w \sum_{l=l_i}^{l_i+q-1} \sum_{k=l_j}^{l_j+q-1} \|\text{Cov}(\mathbf{A}_l, \mathbf{A}_k)\| \\ &\leq cn^{-1} wq \sum_{t=p}^{\infty} \alpha^{\frac{\delta}{2+\delta}} (t) = o(1). \end{aligned}$$

现在证式 (7.5) 中的第二个式子.

$$\text{E}\|S'''_n\|^2 = \sum_{k=w(p+q)+1}^n \text{E}\|\mathbf{V}_k\|^2 + 2 \sum_{w(p+q)+1 \leq i < j \leq n} \|\text{Cov}(\mathbf{V}_i, \mathbf{V}_j)\|.$$

对 $E\|S_n'''\|^2$ 表达式中的第一项, 由 $E\|\mathbf{A}_i\|^2 = E\|\boldsymbol{\eta}_i\varepsilon_i\|^2 \leq E\|\boldsymbol{\eta}_1\|^2\sigma^2 < \infty$ 知

$$\sum_{k=w(p+q)+1}^n E\|\mathbf{V}_k\|^2 \leq \frac{[n-w(p+q)]}{n} E\|\mathbf{A}_k\|^2 \leq c \frac{[n-w(p+q)]}{n} \sigma^2 \leq c \frac{p}{n} = o(1).$$

对 $E\|S_n'''\|^2$ 表达式中的第 2 项有

$$\begin{aligned} & \left\| \sum_{w(p+q)+1 \leq i < j \leq n} \text{Cov}(\mathbf{V}_i, \mathbf{V}_j) \right\| \\ & \leq \frac{1}{n} \sum_{w(p+q)+1 \leq i < j \leq n} \|\text{Cov}(\mathbf{A}_i, \mathbf{A}_j)\| \leq c \frac{[n-w(p+q)]}{n} \sum_{t=1}^{\infty} \alpha^{\frac{\delta}{2+\delta}}(t) \\ & \leq c \frac{p}{n} = o(1) \end{aligned}$$

成立, 所以, $E(S_n''')^2 \rightarrow 0$.

现在证式 (7.6). 通过将 S_n' 中的项表示出来可知,

$$\begin{aligned} \text{Var}(S_n') &= \sum_{m=1}^w \text{Var}(\mathbf{V}_{mn}) + 2 \sum_{1 \leq i < j \leq w} \text{Cov}(\mathbf{V}_{in}, \mathbf{V}_{jn}) \\ &= \sum_{m=1}^w \sum_{i=k_m}^{k_m+p-1} \text{Var}(\mathbf{V}_i) + 2 \sum_{m=1}^w \sum_{k_m \leq i < j \leq k_m+p-1} \text{Cov}(\mathbf{V}_i, \mathbf{V}_j) \\ &\quad + 2 \sum_{1 \leq i < j \leq w} \sum_{i=k_i}^{k_i+p-1} \sum_{j=k_j}^{k_j+p-1} \text{Cov}(\mathbf{V}_i, \mathbf{V}_j) \\ &= \sum_{i=1}^n \text{Var}(\mathbf{V}_i) - \sum_{m=1}^w \sum_{j=l_m}^{l_m+q-1} \text{Var}(\mathbf{V}_j) - \sum_{k=w(p+q)+1}^n \text{Var}(\mathbf{V}_k) \\ &\quad + 2 \sum_{m=1}^w \sum_{k_m \leq i < j \leq k_m+p-1} \text{Cov}(\mathbf{V}_i, \mathbf{V}_j) \\ &\quad + 2 \sum_{1 \leq i < j \leq w} \sum_{i=k_i}^{k_i+p-1} \sum_{j=k_j}^{k_j+p-1} \text{Cov}(\mathbf{V}_i, \mathbf{V}_j). \end{aligned} \tag{7.9}$$

对式 (7.9) 中的各项进行分析得,

$$\text{Var}(S_n') = \text{Var}\left(\sum_{i=1}^n \mathbf{V}_i\right) + o(1).$$

具体分析过程如下.

$$\begin{aligned} & \sum_{1 \leq i < j \leq w} \sum_{i=k_i}^{k_i+p-1} \sum_{j=k_j}^{k_j+p-1} \|\text{Cov}(\mathbf{V}_i, \mathbf{V}_j)\| \\ &= n^{-1} \sum_{1 \leq i < j \leq w} \sum_{i=k_i}^{k_i+p-1} \sum_{j=k_j}^{k_j+p-1} \|\text{Cov}(\mathbf{A}_i, \mathbf{A}_j)\| \end{aligned}$$

$$\begin{aligned} &\leq cn^{-1}w \sum_{i=k_i}^{k_i+p-1} \sum_{j=k_j}^{k_j+p-1} \alpha^{\frac{\delta}{2+\delta}}(j-i) \leq cn^{-1}wp \sum_{t=q}^{\infty} \alpha^{\frac{\delta}{2+\delta}}(t) \\ &\leq cn^{-1}wpq^{1-\frac{\delta\gamma}{2+\delta}} \leq cq^{1-\frac{\delta\gamma}{2+\delta}} = o(1), \end{aligned} \quad (7.10)$$

$$\sum_{m=1}^w \sum_{j=l_m}^{l_m+q-1} \text{Var}(\mathbf{V}_j) \leq n^{-1} \sum_{m=1}^w \sum_{j=l_m}^{l_m+q-1} E\|\mathbf{A}_i\|^2 \leq cn^{-1}wq\sigma^2 = o(1), \quad (7.11)$$

$$\sum_{i=w(p+q)+1}^n \text{Var}(\mathbf{V}_i) = n^{-1} \sum_{i=w(p+q)+1}^n \text{Var}(\mathbf{A}_i) \leq cn^{-1}p = o(1). \quad (7.12)$$

注意到

$$\begin{aligned} \sum_{1 \leq i < j \leq n} \text{Cov}(\mathbf{V}_i, \mathbf{V}_j) &= \sum_{m=1}^w \sum_{i=k_m}^{k_m+p-1} \sum_{i < j \leq n} \text{Cov}(\mathbf{V}_i, \mathbf{V}_j) + \sum_{m=1}^w \sum_{i=l_m}^{l_m+q-1} \sum_{i < j \leq n} \text{Cov}(\mathbf{V}_i, \mathbf{V}_j) \\ &\quad + \sum_{k=w(p+q)+1}^n \sum_{k < j \leq n} \text{Cov}(\mathbf{V}_k, \mathbf{V}_j) \\ &:= R_1 + R_2 + R_3. \end{aligned} \quad (7.13)$$

可证得

$$R_1 = \sum_{m=1}^w \sum_{k_m \leq i < j \leq k_m+p-1} \text{Cov}(\mathbf{V}_i, \mathbf{V}_j) + o(1), \quad R_2 = o(1), \quad R_3 = o(1).$$

首先, 考虑项 R_1 .

$$\begin{aligned} R_1 &= \sum_{m=1}^w \sum_{i=k_m}^{k_m+p-1} \sum_{i < j \leq k_m+p-1} \text{Cov}(\mathbf{V}_i, \mathbf{V}_j) + \sum_{m=1}^w \sum_{i=k_m}^{k_m+p-1} \sum_{s=m}^w \sum_{j=l_s}^{l_s+q-1} \text{Cov}(\mathbf{V}_i, \mathbf{V}_j) \\ &\quad + \sum_{m=1}^w \sum_{i=k_m}^{k_m+p-1} \sum_{j=w(p+q)+1}^n \text{Cov}(\mathbf{V}_i, \mathbf{V}_j) \\ &:= R_{11} + R_{12} + R_{13}. \end{aligned}$$

对于 R_{12}, R_{13} , 有

$$\begin{aligned} \|R_{12}\| &\leq n^{-1} \sum_{i=1}^n \sum_{m=1}^w \sum_{j=l_m}^{l_m+q-1} \|\text{Cov}(\mathbf{A}_i, \mathbf{A}_j)\| \\ &\leq cn^{-1}wq \sum_{t=1}^{\infty} \alpha^{\frac{\delta}{2+\delta}}(t) \leq c \frac{q}{p} = o(1), \\ \|R_{13}\| &= n^{-1} \sum_{m=1}^w \sum_{i=k_m}^{k_m+p-1} \sum_{j=w(p+q)+1}^n \|\text{Cov}(\mathbf{A}_i, \mathbf{A}_j)\| \\ &\leq n^{-1} \sum_{m=1}^w \sum_{i=k_m}^{k_m+p-1} \sum_{j=i-q}^{\infty} \alpha^{\frac{\delta}{2+\delta}}(j-i) \end{aligned}$$

$$\leq n^{-1}wp \sum_{t=q}^{\infty} \alpha^{\frac{\delta}{2+\delta}}(t) \longrightarrow 0, \quad n \longrightarrow \infty.$$

所以,

$$R_1 = \sum_{m=1}^w \sum_{k_m \leq i < j \leq k_m + p - 1} \text{Cov}(\mathbf{V}_i, \mathbf{V}_j) + o(1). \quad (7.14)$$

类似于 R_{13} ,

$$\begin{aligned} R_2 &= n^{-1} \sum_{m=1}^w \sum_{j=l_m}^{l_m+q-1} \sum_{j < l \leq n} \text{Cov}(\mathbf{A}_j, \mathbf{A}_l) \\ &\leq n^{-1} \sum_{l=1}^n \sum_{m=1}^w \sum_{j=l_m}^{l_m+q-1} \text{Cov}(\mathbf{A}_j, \mathbf{A}_l) \\ &= o(1) \end{aligned} \quad (7.15)$$

及

$$\begin{aligned} R_3 &= n^{-1} \sum_{k=w(p+q)+1}^n \sum_{j=k+1}^n \text{Cov}(\mathbf{A}_k, \mathbf{A}_j) \\ &\leq cn^{-1}[n - w(p+q)] \sum_{j-k=1}^{\infty} \alpha^{\frac{\delta}{2+\delta}}(j-k) \\ &\leq cn^{-1}p \sum_{t=1}^{\infty} \alpha^{\frac{\delta}{2+\delta}}(t) \leq cn^{-1}p \rightarrow 0, \quad n \longrightarrow \infty. \end{aligned} \quad (7.16)$$

把式 (7.10)–(7.16) 代入式 (7.9) 中有

$$\begin{aligned} \text{Var}(S'_n) &= \sum_{i=1}^n \text{Var}(\mathbf{V}_i) + 2 \sum_{1 \leq i < j \leq n} \text{Cov}(\mathbf{V}_i, \mathbf{V}_j) + o(1) \\ &= \text{Var}\left(\sum_{i=1}^n \mathbf{V}_i\right) + o(1) = \Sigma + o(1), \end{aligned}$$

故式 (7.6) 成立.

由引理 7.1 知,

$$\left\| \text{Eexp}\left(it \sum_{m=1}^w \mathbf{V}_{mn}\right) - \prod_{m=1}^w \text{Eexp}(it\mathbf{V}_{mn}) \right\| \leq 16w\alpha(q+1) \leq c \frac{n\alpha(q)}{p} = o(1).$$

故式 (7.7) 成立.

最后证明式 (7.8) 成立. 假设对 $\varepsilon > 0$, 有 $\lambda > 2\varepsilon > 0$. 根据引理 7.3 可知

$$\sum_{m=1}^w \text{E}\|\mathbf{V}_{mn}\|^2 I(\|\mathbf{V}_{mn}\| > \varepsilon)$$

$$\begin{aligned}
&\leq n^{-2-\lambda} \sum_{m=1}^w \mathbb{E} \left\| \sum_{k=k_m}^{k_m+p-1} \mathbf{A}_k \right\|^{2+\lambda} \\
&\leq n^{-2-\lambda} w \left(p^{1+\lambda/2} \max_{1 \leq i \leq n} \|\mathbf{A}_i\|_{2+\delta}^{2+\lambda} + p^{1+\varepsilon} \max_{1 \leq i \leq n} \|\mathbf{A}_i\|_{2+\delta}^{2+\lambda} \right) \\
&\leq cn^{-2-\lambda} wp^{1+\lambda/2} \leq cn^{-1-\lambda} wp^{\lambda/2} \rightarrow 0, \quad n \rightarrow \infty.
\end{aligned}$$

综上所述式 (7.4) 成立. 另一方面, 注意到

$$\Sigma = \text{Var} \left(\frac{1}{n^{1/2}} \sum_{i=1}^n \boldsymbol{\eta}_i \varepsilon_i \right) = \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^n \text{Cov}(\boldsymbol{\eta}_i \varepsilon_i, \boldsymbol{\eta}_j \varepsilon_j) = \frac{\mathbf{B}}{n} \sum_{i=1}^n \mathbb{E}(\varepsilon_i \varepsilon_i) = \mathbf{B} \sigma^2. \quad (7.17)$$

由式 (7.17) 及式 (7.1)–(7.4) 知定理 3.1 的结论成立.

定理 3.2 的证明 基于引理 7.4, 此定理的证明类似于 [10] 中定理 3.2 的证明, 故此省略. 证毕.

参 考 文 献

- [1] Ramsa J O, Silverman B W. Functional Data Analysis. New York: Springer-Verlag, 1997
- [2] Cardot H, Ferraty F, Sarda P. Functional linear model. *Statistics & Probability Letters*, 1999, 45: 11–22
- [3] Cardot H, Ferraty F, Sarda P. Spline estimators for the functional linear model. *Statistica Sinica*, 2003, 13: 571–591
- [4] Cai T T, Hall P. Prediction in functional linear regression. *Annals of Statistics*, 2006, 34: 2159–2179
- [5] Hall P, Horowitz J L. Methodology and convergence rates for functional linear regression. *Annals of Statistics*, 2007, 35: 70–91
- [6] Li Y, Hsing T. On rates of convergence in functional linear regression. *Journal of Multivariate Analysis*, 2007, 98: 1782–1804
- [7] Crambes C, Kneip A, Sarda P. Smoothing splines estimators for functional linear regression. *Annals of Statistics*, 2009, 37: 35–72
- [8] Aneiros-Pérez G, Vieu P. Semi-functional partial linear regression. *Statistics & Probability Letters*, 2006, 76: 1102–1110
- [9] Aneiros-Pérez G, Vieu P. Nonparametric time series prediction: a Semi-functional partial linear modeling. *Journal of Multivariate Analysis*, 2008, 99: 834–857
- [10] Shin H. Partial functional linear regression. *Journal of statistical planning and inference*, 2009, 139: 3405–3418
- [11] Shin H, Lee M. On prediction rate in partial functional linear regression. *Journal of Multivariate Analysis*, 2012, 103: 93–106
- [12] Fan G L, Liang H Y, Wang J F, Xu H X. Asymptotic properties for LS estimators in EV regression model with dependent errors. *Advances in Statistical Analysis*, 2010, 94: 89–103

- [13] Wiecezorek B, Ziegler K. On optimal estimation of a non-smooth mode in a nonparametric regression model with α -mixing errors. *Journal of Statistical Planning and Inference*, 2010, 140(2): 406–418
- [14] Fan G L, Liang H Y. Empirical likelihood for longitudinal partially linear model with α -mixing errors. *Journal of Systems Science and Complexity*, 2013, 26(2): 232–248
- [15] Riesz F, Sz-Nagy B. *Functional Analysis*. New York: Dover Publications, 1955
- [16] Lv Y, Du J, Sun Z M. Functional partially linear quantile regression model. *Metrika*, 2014, 77: 317–332
- [17] Volkonskii V A, Rozanov Y A. Some limit theorems for random functions. *Theory of Probability & Its Applications*, 1959, 4: 178–197
- [18] Hall P, Heyde C C. *Martingale Limit Theory and Its Application*. New York: Academic Press, 1980
- [19] Shao Q M, Yu H. Weak convergence for weighted empirical processes of dependent sequences. *Annals of Probability*, 1996, 24: 2098–2127

Partial Functional Linear Models with Dependent Errors

WANG YAFEI

(College of Applied Sciences, Beijing University of Technology, Beijing 100124, China)

(E-mail: wangyafei@emails.bjut.edu.cn)

DU JIANG ZHANG ZHONGZHAN

(College of Applied Sciences, Beijing University of Technology,

Collaborative Innovation Center on Capital Social Construction

and Social Management, Beijing 100124, China)

(E-mail: dujiang@bjut.edu.cn; z Zhang@bjut.edu.cn)

Abstract In this paper, we study the estimation of partial functional linear regression models with stationary α -mixing random error sequence. With approximating to the slope function by the Karhunen-Loève expansion, we propose an estimation method for the unknown parameters and the slope function. The asymptotic normality of the proposed parameter estimators and the convergence rate of the slope function estimator are established. Intensive simulation experiments and real data analysis are conducted to show that the proposed method performs well with a finite sample.

Key words asymptotic normality; α -mixing sequence; principal components; convergence rate

MR(2000) Subject Classification 46B20

Chinese Library Classification O177.2